Automated Planning for Hydrothermal Vent Prospecting Using AUVs: RSMG Report 8

Zeyn A Saigol

Thesis group members	Richard Dearden (supervisor)
	Jeremy Wyatt (supervisor)
	Xin Yao (RSMG member)
	Hamid Dehghani (non-RSMG member)

School of Computer Science
College of Engineering and Physical Sciences
The University of Birmingham
28th April 2010

Contents

1 Introduction

The aim of my PhD is to develop novel AI techniques that will enable Autonomous Underwater Vehicles (AUVs) to locate hydrothermal vents, which are superheated outgassings of water found on the ocean floor. Vents emit a chemically-altered plume that can be detected from dozens of kilometres away, but they are still hard to find for two reasons: firstly there is a vast volume of ocean to search, and secondly even once a plume is detected it is hard to trace it back to the source, due to factors such as turbulent mixing. Current methods for finding hydrothermal vents rely on manually defining an area for the AUV to perform an exhaustive search over (or on using remotely-operated vehicles instead of AUVs). My work has developed planning algorithms that should allow an AUV to use on-board data processing and decision making to find vents more efficiently, maximising the returns for missions which are limited in duration by battery power.

I decided to split the problem into two parts: firstly finding an estimate of the location of nearby vents based on the observation history, and secondly finding a plan based on this probabilistic map of vent locations. I adopted Michael Jakuba's algorithms (Jakuba, 2007) to solve the problem of creating a map of hydrothermal vents using AUV sensor data. Jakuba's work is based on the occupancy grid (OG) method (Martin and Moravec, 1996; Elfes, 1989), where the search area is divided up into a discrete grid. Using Jakuba's approach (and code he very kindly provided me with) allowed me to concentrate on the planning side of the problem.

I formulated the planning problem as a partially-observable Markov decision process (POMDP), where the system state (which cannot be directly observed by the agent) consists of the location of the AUV, the resources remaining to the AUV, and the location(s) of nearby vents. The only action available is setting the vehicle's heading, and the transition function gives the probability distribution over resulting locations. The observations are measurements of various biogeochemical 'tracers' produced by vents, and the observation function essentially models how tracer concentration varies with distance from a vent. Finally, the reward function rewards the agent only when a new vent is found.

While there are now fairly efficient algorithms available for solving POMDPs, the continuous state and action spaces, together with limited resources available to the agent, mean this problem is completely intractable for them. I have developed and implemented several algorithms that produce approximate solutions to the POMDP, which I have termed *information-lookahead* (IL) algorithms. The first step in developing these algorithms was to create an *information state Markov decision process* (ISMDP) model based on the POMDP problem formulation. This required me to derive an observation function from the environmental model underpinning Jakuba's OG algorithm.

The IL algorithms solve an ISMDP approximately, and belong to the family of online POMDP solvers (Ross et al., 2008). They work by simulating all possible actions and subsequent observations several steps into the future, and using these simulations to calculate the value of each action. Computational considerations mean that simulations can only be run a small number of steps into the future, so a heuristic value has to be chosen for 'leaf' belief states. I developed a *certainty-equivalent* (CE) heuristic that just assumes the OG found in a leaf state is correct, and solves the resulting MDP. The algorithms themselves and experiments comparing the different algorithms are documented in an IJCAI paper I co-authored (Saigol et al., 2009).

Subsequently I have switched to applying entropy-based methods to this domain. In the context of trying to find item(s) of interest, entropy can only be used as a heuristic, and the algorithms I created do not directly approximate the POMDP solution. Since my last progress report (Saigol, 2009), I have:

- Developed a family of entropy-based algorithms, and an orienteering-problem solver which can be used to enhance both the IL and new belief-change-maximisation algorithms (described in Section 2).
- Performed experiments with these new belief-change-maximisation algorithms and written a paper describing the algorithms and experiments.
- Presented a talk at the IRLab meeting on 2nd February 2010, titled *Using entropy to drive search in occupancy grids.*
- Started writing my thesis (see Section 3).

2 Technical Progress

This section describes the new algorithms I have developed since my last progress report (Saigol, 2009). I also performed a series of experiments to evaluate these algorithms, using slightly different experimental conditions and more samples than the previous IL experiments. The algorithms and experiments are described in more detail in Saigol et al. (2010) (submitted).

2.1 $\Sigma\Delta H$ Algorithms

The $\Sigma\Delta H$ algorithms were inspired by entropy reduction methods that are commonly used for robot mapping (e.g. Bourgault et al., 2002; Stachniss et al., 2005). The idea of these is that the robot takes the action that will reduce the entropy of the map most; however in my domain, the low prior probability of a cell in the OG being occupied means that the entropy of the map usually *increases* when the agent detects a plume. As plume detections are useful, I instead chose to maximise the change in entropy, given by the sum (over all cells c) of the absolute value of the entropy change for each cell, H_c. The basic $\Sigma\Delta H$ algorithm is then for the agent to choose the action that maximises

$$\mathrm{E}_z\left[\sum_c \left|H_c(b') - H_c(b)\right|\right] \tag{1}$$

where $\mathrm{E}_z[\ldots]$ represents the expectation over possible observations z, and b' is the belief state that results from starting in b, performing the action under consideration, and observing z.

A further refinement, resulting in significantly improved vent-hunting performance, is to convert the single-action lookahead formulation above into a discounted infinite-horizon MDP. The $\Sigma\Delta H$-MDP algorithm effectively assumes that the agent can move anywhere instantly, by calculating $\Sigma\Delta H$ values (Equation 1) for all cells in the grid, and then solving an MDP using these values. The intuition behind the approach is that it ignores observations gained from future actions, but uses the expected value of making an observation at a cell (given the current belief state) as the value of the cell.

2.2 OP Correction

The $\Sigma\Delta H$-MDP algorithm implicitly assumes that the same value can be gained by making an observation from the same cell twice. In reality this is not the case, for example if a plume is detected at a cell, further measurements at that cell will make almost no difference to the occupancy grid. To address this issue I converted the MDP over $\Sigma\Delta H$ values into an orienteering problem (Golden et al., 1987; Tsiligirides, 1984), a variant of the travelling salesman problem. The orienteering problem (OP) defines a score for each city, and the aim is to find a route over a subset of cities to maximise the total score while respecting a maximum total distance. For an occupancy grid, each cell is a city, and the main challenge for implementation is the limited connectivity between cities. I implemented a Monte-Carlo OP solver that samples random paths to find the best. As paths that cross themselves cannot be optimal, each path sampled by my implementation is a Self-Avoiding Walk (Sokal, 1995), which is costly to generate.

The $\Sigma\Delta H$-OP algorithm produced statistically significant improvements over $\Sigma\Delta H$-MDP. The OP correction was also applied to the IL approach, as a replacement for CE as the leaf-node heuristic. In this case, while IL-OP worked much better than IL-CE, using OP lookahead did not conclusively improve performance compared with using the basic leaf-node heuristic (i.e. the occupancy probability of the leaf cell).

3 Thesis Progress

3.1 Thesis Outline

1. Introduction
    - 1.1. Motivation
    - 1.2. Problem Overview
    - 1.3. Evaluation
    - 1.4. Contributions
        - 1.4.1. Formal model of the vent prospecting problem
        - 1.4.2. Entropy-inspired planning methods
        - 1.4.3. Information-lookahead planning methods
        - 1.4.4. Orienteering problem correction
    - 1.5. Document Structure
2. Oceanography Background
    - 2.1. Hydrothermal Vents
    - 2.2. Vent Prospecting
        - 2.2.1. Chemical and Physical Tracers
    - 2.3. Work with ABE

   7.1. MDP Issue and OP Solution
   7.2. Application to IL Algorithms
   7.3. Results and Discussion

8. Conclusions and Future Work

   8.1. Summary
   8.2. Evaluation
   8.3. Conclusions
   8.4. Future Work

A Glossary of Terms

B List of Symbols

3.2 Progress By Chapter

- Chapters 1 (Introduction), 2 (Oceanography Background), and 3 (Problem Model) – draft completed.
- Chapter 4 (AI Background) – a lot of the material for this chapter can be taken from my thesis proposal, the IJCAI paper and the belief-change-maximisation paper. However, this material needs to be re-organised, and there are several additional areas that need to be covered (particularly robotic exploration using entropy (e.g. Bourgault et al., 2002; Stachniss et al., 2005; Low et al., 2009) and adaptive optimal sensing work, including submodularity (Singh et al., 2009)).
- Chapters 5 (Planning Using Entropy) and 7 (Using Orienteering Methods as a Correction) – will be based heavily on the contents of the belief-change-maximisation paper, with a few additional details.
- Chapter 6 (Planning Using Information Lookahead) – will be based heavily on the contents of the IJCAI paper, with a few additional details.
- Chapter 8 (Conclusions and Future Work) – to be written.

4 Future Plans

4.1 Research Tasks

While not a priority, I would like to conduct further research to understand the behaviour of my algorithms better. In particular, I would like to investigate:

- Why the OP correction does not improve the performance of IL (for example, IL4-OP6 is no better than IL4).
- What aspects of the observation model and/or occupancy grid algorithm allow IL to produce very good performance even using just four steps of lookahead.

Table 1: Timetable for writing up my thesis

Task	*Duration (days)*	*Start Date*	*End Date*
Ch4: AI background	15	May-03	May-25
Ch5: Entropy	10	May-26	Jun-15
Ch6: Info-Lookahead	15	Jun-16	Jul-06
Ch7: OP correction	15	Jul-07	Jul-27
Ch8: Conclusions	10	Jul-28	Aug-10
Re-writing (esp Intro)	10	Aug-11	Aug-24
LaTeX formatting and printing	5	Aug-25	Aug-31
Planned submission day	1	Sep-01	Sep-01

4.2 Timetable

The revised timetable for my PhD is shown in Table 1. I started writing up on 8^{th} March, my target submission date is Wednesday 1^{st} September, and my submission deadline is 1^{st} October 2010.

Currently I am a self-funded student.

References

Bourgault, F., Makarenko, A., Williams, S., Grocholski, B., and Durrant-Whyte, F. (2002). Information based adaptive robotic exploration. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*.

Elfes, A. (1989). Using occupancy grids for mobile robot perception and navigation. *Computer*, 22(6):46–57.

Golden, B. L., Levy, L., and Vohra, R. (1987). The orienteering problem. *Naval Research Logistics*, 34(3):307–318.

Jakuba, M. (2007). *Stochastic Mapping for Chemical Plume Source Localization with Application to Autonomous Hydrothermal Vent Discovery*. PhD thesis, MIT and WHOI Joint Program.

Low, K. H., Dolany, J. M., and Khosla, P. (2009). Information-theoretic approach to efficient adaptive path planning for mobile robotic environmental sensing. In Alfonso Gerevini, Adele Howe, A. C. and Refanidis, I., editors, *Proceedings of the Nineteenth International Conference on Automated Planning and Scheduling (ICAPS-09)*, pages 233–240. The AAAI Press, Menlo Park, California.

Martin, M. C. and Moravec, H. (1996). Robot evidence grids. Technical Report CMU-RI-TR-96-06, CMU Robotics Institute.

Ross, S., Pineau, J., Paquet, S., and Chaib-draa, B. (2008). Online planning algorithms for POMDPs. *Journal of Artificial Intelligence Research*, 32:663–704.

Saigol, Z. A. (2009). Intelligent planning for Autonomous Underwater Vehicles: RSMG report 7. School of Computer Science, University of Birmingham.

Saigol, Z. A., Dearden, R. W., Wyatt, J. L., and Murton, B. J. (2009). Information-lookahead planning for AUV mapping. In *Proceedings of the Twenty-first International Joint Conference on Artificial Intelligence (IJCAI-09)*.

Saigol, Z. A., Dearden, R. W., Wyatt, J. L., and Murton, B. J. (2010). Belief change maximisation for hydrothermal vent hunting using occupancy grids. In *(submitted)*.

Singh, A., Krause, A., and Kaiser, W. J. (2009). Nonmyopic adaptive informative path planning for multiple robots. In *Proceedings of the 21st international jont conference on Artifical intelligence (IJCAI-09)*, pages 1843–1850, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Sokal, A. D. (1995). *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, chapter Monte Carlo Methods for the Self-Avoiding Walk. Oxford University Press, New York. arXiv:hep-lat/9405016.

Stachniss, C., Grisetti, G., and Burgard, W. (2005). Information gain-based exploration using Rao-Blackwellized particle filters. In *Proc. of Robotics, Science and Systems*, pages 65–72.

Tsiligirides, T. (1984). Heuristic methods applied to orienteering. *Journal of the Operational Research Society*, 35(9):797–809.